Cost Estimation in Heterogeneous Cloud Environment using Fault Tolerance Services

[1]K. Chitra, [2]Dr. Sivaprakasam,
[1]Research Scholar, Mother Teresa Women's University, Kodaikanal, Tamilnadu INDIA
[2]Associate Professor, Sri Vasavi College, Erode, Tamlinadu, INDIA

***Abstract*—Clouds have emerged as a computing infrastructure that enables rapid delivery of computing resources as utility in a dynamically scalable, Virtualized manner. Cloud computing provides virtually unlimited computational resources to its users, while letting them pay only for the resources they actually use at any given time. Cloud service providers offer resources to users on demand using various pricing schemes based on on-demand instances, reserved instances and spot instances. A data center in the cloud is a large distributed computing environment consisting of heterogeneous resources. The proposed strategy improves the resource provisioning mechanisms by introducing the E-Auction in which we provide the optimal bidding mechanism by considering the workload selection in terms of jobs deadline, average CPU time consumption, job submission and job remaining time. Also, in order to optimize or minimize the cost within the performance requirement, we present the workflow optimization logics.**

***Index Terms*— Cloud Computing, Consumer Expected Cost, Provider Expected Cost, Job Migration, Fault Tolerant Services, Virtual Machine.**

I. INTRODUCTION

Cloud computing has become an important computing infrastructure for many applications. Provisioning, allocating, and pricing the VM instances are important issues that have to be maintained by cloud providers. We apply the transformation logics across multiple cloud VMs, transformation logic that minimizes the resource provisioning cost. Transformations logics consist of effective scheduling mechanisms and the maximum revenue. Some cloud jobs are having the dependence between the jobs. At that time, we can't able to predict the optimal VM with dependency in allocation of resources. To overcome the existing problem, a novel e-auction model is proposed for provisioning the resources. In our proposed system, we concentrate with the energy and load distribution of the data in the e-auction model for provisioning the resource. Energy models are used to consume energy at different level (queue energy, delay energy, consumption energy). Then, we focus on the fault tolerance strategies based on the migration of the auction resources in the double auction manner. We implement fault tolerate services with migration of jobs in the distributed cloud environment. From this proposal, we can able to achieve minimum cost, less fault value, easy to recover jobs, reduce the time to execute job, workflow execution, load distribution.

The fault tolerance services are used in the cloud with job Migration. In this approach, Nash Auction Equilibrium (NAE) algorithm is used to check whether the PEC (Provider Expected Cost) value is less than or equal to the CEC (Consumer Expected Cost) value. The PEC and CEC values related calculations are available in section-III. Once the VM has been found and the job in process, job may fail at any time during process. So the job migration is needed at this time of failure. That failed job is reallocated to other available VM along with the optimal cost and time. Double auction is required in this situation. The following steps to be performed for fault tolerance services. (i) Several VMs are created over the heterogeneous cloud environment. (ii) Each and every VM can be allocated through the cloudlets. (iii) Compute Consumer Expected cost (CEC) for all jobs. (iv) Estimate Provider Expected Cost (PEC) for virtual machines. (v) Match CEC and PEC values using Nash Auction Equilibrium Algorithm.(vii) Start simulation for utilization (vii) After verifying the utilization status, then the migration frequency will be evaluated.

The section II shows the related work based on the concept of job migration, fault tolerance services, work flows, job reallocation. The section III deals with the calculations for CEC and PEC values. Section IV deals with the comparison of proposed system with the already existing approaches. Section V describes the conclusion of this paper and future work related to this paper. Section VI deals with the various reference papers available in various journals.

II. RELATED WORK

Gideon Juve, Ewa Deelman [1] have discussed some of the resource provisioning challenges that affect workflow applications, and They described how provisioning techniques such as advance reservations and multi level scheduling can improve workflow performance by minimizing or eliminating these workflow overheads. They also discussed how emerging Iaas platforms can be combined with multi level scheduling to create flexible and efficient execution environments for workflow applications. Mousumi Paul, Debabrata Samanta, Goutam Sanyal [2] established a scheduling mechanism which

follows the lexi search approach to assign the tasks to the available resources.The scheduled task will be maintained by load balancing algorithm that distribute the pool of task into small partition and then distribute into local middleware. Raj Kumar Buyya, Anton Beloglazov and Jemal Abawajy [3] worked together to advance the cloud computing field in two ways. In that two way approach, first one plays a significant role in the reduction of data center energy consumption costs and helps to develop a strong competitive cloud computing industry. Second approach deals with consumers become increasingly conscious about the environment. Constanino Vazquez, Eduardo Huedo, Ruben.S.Montero [4] have proposed and evaluated architecture to build a grid infrastructure in a very flexible way, being able to hire computer resources with potentially different interfaces. K.C.Gowda, Radhika TV, Akshatha .M [5] presented the method for efficient resource allocation that will help cloud owner to reduce wastage of resources and to achieve maximum profit. They have described a work on the allocation of resources in a dynamic cloud environment by using priority algorithm which decides the allocation sequence for different jobs requested among the different users after considering the priority based on some optimum threshold decided by the cloud owner. Stefan Schulte, Diecter Schuller, Philipp Hoenisch [6] presented the Vienna platform for elastic processes which combines the functionalities of BPMS with that of a cloud resource management system. They have also presented an extended optimization model and heuristic for workflow scheduling and resource allocation for elastic process execution.

Christina Hoffa , Gaurang Mehta,Timothy Freeman [7] aimed to quantify the performance differences in an application, an application scientist would see while running workflows on various types of resources including both physical and virtual environments. Wenbing Zhao, D.M .Melliar-smith and L.E.Moser [8] explained about Low Latency Fault Tolerance (LLTF) middleware, that provides fault tolerance for distributed applications deployed within an cloud computing or data center environment, The LLFT protocol provides a reliable, totally ordered multicast service by communicating message ordering information from the primary in a group to the backups in the group. Md Imran Alam, Manjusha Pandey, Siddharth .S. Rautaray [9] designed Tabu search algorithm for finding quality solution in less time with less resources. They considered three main factors. In this algorithm, those are optimum solution, dead line constraint and cost constraint. Anju Bala, Inderveer chana [10] discussed the fault tolerance techniques covering its research challenges, tools used for implementing fault tolerance techniques in cloud computing. Cloud virtualized system architecture is also proposed based on HAProxy. Autonomic fault tolerance is implemented dealing with various software faults in cloud virtualized server application environments.

Awada Uchechukwu , Keqiu Li ,Yanming Shen [11] have presented energy consumption formulas for calculating the total energy consumption in cloud environments and show that there are incentives to save energy. They described an energy consumption tools and empirical analysis approaches. They provided generic energy consumption models for server idle and server active states. Jing Liu, Xing-Guo Luo [12] have established a scheduling model for cloud computing based on MO-GA algorithm to minimize energy consumption and maximize the profit of service providers under the constraints of deadlines. They first proposed a job scheduling architecture under the environment of cloud computing which contains several components to analyze the applications to improve the effectiveness and efficiency of the computing. Nidhi Jain Kansal and Inderveer Chana [13] have presented a systematic review of existing load balancing techniques. Their study concludes that all the existing techniques mainly focus on reducing associated overhead service response time and improving performance of various parameters are identified and they are used to compare the existing techniques.

M.Sudha, M.Monica [14] investigated the efficient management of workflows in cloud computing. The result of investigation show that a workflow with short job runtimes, the virtual environment can provide good compute time performance but it can suffer from resource scheduling delays and wide area communications. Marcos Dias de Assuncao et al ., [15] have evaluated the cost of improving the scheduling performance of virtual machine requests by allocating additional resources from a cloud computing infrastructure. Haiyang Qian and Deep Medhi [16] used the wear and tear cost and power consumption to capture the server operational cost in CSP data centers. Abirami .S.P and Shalini Ramanathan [17] developed effective scheduling algorithm named LSTR scheduling algorithm which schedules both the task and resources are designed. This algorithm mainly focuses in eradicating the starvation and deadlock conditions. Mladen A.Vouk [18] discussed the concept of cloud computing and try to address some of issues, related research topics and cloud implementations available today. Amelie chi Zhou and Bingsheng [19] proposed a workflow transformation base optimization framework namely ToF. They formulate the performance and cost optimizations of workflows in the cloud as transformation and optimization. They have designed two components to be extensible for user requirements on performance and cost, cloud offerings and workflows. Martin Bichler, Jayant Kalagnanam [20] have discussed a number of winner determination problems in the context of multi-attribute auctions.

III. COST ESTIMATION CALCULATION

A. CONSUMER EXPECTED COST COMPUTATION (CEC computation)

The consumer expected cost value is calculated by using the following parameters (i) cost (ii) time (iii) Energy (iv) Demand (v) Supply.

TABLE-I

Description of parameters used in the CEC Computation

c_i Cost for i[th task]
v_i Value for i[th] task
ec_i Execution cost for i[th] task
er_i Energy Rate for i[th] task
nc_i Energy cost for i[th] VM
dr_i Demand Rate for i[th] task
σ_i Demand for i[th] task
β_i Supply for i[th] task
S_i Size of i[th] task
τ_i Time Rate for i[th] task
st_i Submission Time of i[th] task
rt_i Required time of i[th] task
ϑ_i Required energy of i[th] task

The above table explains the parameters used in consumer expected cost computation. This CEC Value is calculated by adding cost for i[th] job and value for i[th] job and the result is divided by 2.

$$\text{CEC} = \frac{(c_i + v_i)}{2} \quad (1)$$

c_i and v_i are calculated as follows

$$c_i = (S_i * ec_i) + (S_i * nc_i) \quad (2)$$

$$v_i = (\tau_i + dr_i + er_i) \quad (3)$$

The Time rate τ_i is calculated by adding the Submission time of i[th] job and required time of i[th] job.

$$\tau_i = (st_i + rt_i) \quad (4)$$

The demand rate is calculated by dividing the demand and supply to the i[th] job.

$$dr_i = \frac{\sigma_i}{\beta_i} \quad (5)$$

The energy rate is equal to the required energy of the i[th] job.

$$er_i = \vartheta_i \quad (6)$$

B. PROVIDER EXPECTED COST COMPUTATION (PEC Computation)

This cost is determined by the server agent. The provider expected cost is calculated by using the following parameters (i) Time (ii) cost (iii) Load (IV) Energy (v) Demand (vi) Supply (vii) Speed

TABLE – II
The Parameters used in calculation of PEC computation

p_j Price for j[th] VM
γ_j Value for j[th] VM
wt_j Execution Time for j[th] VM
ep_j Execution unit Price for j[th] VM
E_j Energy Rate in j[th] VM
qe_j Queuing Energy in j[th] VM
μ_j Energy unit Price for j[th] VM
δ_j Success Rate for j[th] VM
ω_j Demand Rate for j[th] VM
σ_j Load and Energy value
d_j Demand
Sp_j Supply
L Load
NS_j Number of Successive Jobs in j[th] VM
NO_j Number of Overall Jobs in j[th] VM
NF_j Number of Failure Jobs in j[th] VM
ε_{jt} Current Energy for j[th] VM in t[th] time
e_j Overall Energy for j[th] VM
L_{jt} Load for j[th] VM in t[th] time
ρ_j Speed of j[th] VM
∂_i i[th] Job Size

The PEC value is calculated as follows

$$\text{PEC} = \frac{(p_j + \gamma_j)}{2} \quad (7)$$

The price for j[th] VM is calculated by using the following parameters (i) Execution time (ii) execution unit price (iii) energy rate (iv) queuing energy (v) energy unit price

$$p_j = ((wt_j * ep_j) + (1 - E_j + qe_j) * \mu_j) \quad (8)$$

The value for jth VM is calculated by adding success rate, Demand rate and load, energy value divided by 3.

$$\gamma_j = \frac{(\delta_j + \omega_j + \sigma_i)}{3} \quad (9)$$

The demand rate is calculated by dividing demand for j[th] VM and supply for j[th] VM.

$$\omega_j = \frac{d_j}{Sp_j} \quad (10)$$

The success rate is the ratio of number of successive jobs in the j[th] VM to the number of overall jobs in the j[th] VM.

$$\delta_j = \frac{NS_j}{NO_j} \quad (11)$$

The overall jobs in the j[th] VM is the sum of success and failure jobs in j[th] VM.

$$NO_j = NS_j + NF_j \quad (12)$$

Load and Energy value in j[th] VM is calculated as follows

$$\sigma_j = \frac{(L + E_j)}{2} \quad (13)$$

The Energy rate of j[th] VM in the t[th] time is calculated as

$$E_{jt} = \frac{\varepsilon_{jt}}{e_j} \quad (14)$$

The Load value is calculated as follows

$$L = \frac{L_{jt}}{Max(L_j)} \quad (15)$$

Execution time for i[th] VM is calculated as

$$wt_j = \frac{\sum_{i=1}^{n} \partial_i}{\rho_j} \quad (16)$$

IV. PERFORMANCE *ANALYSIS*

This section deals with the performance of execution time, Makespantime, Migration Frequency, Migration time, Energy in cloud environment with respect to job size, optimal cost, no of jobs.

Fig.1

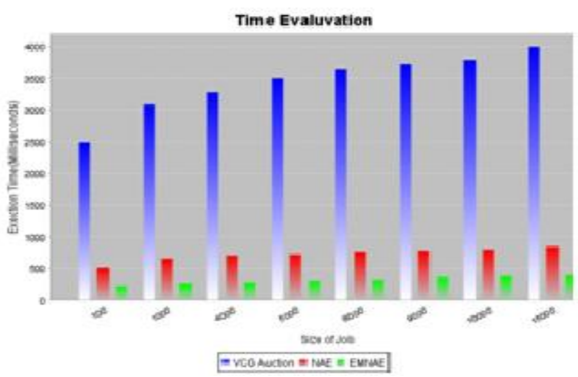


Fig.2

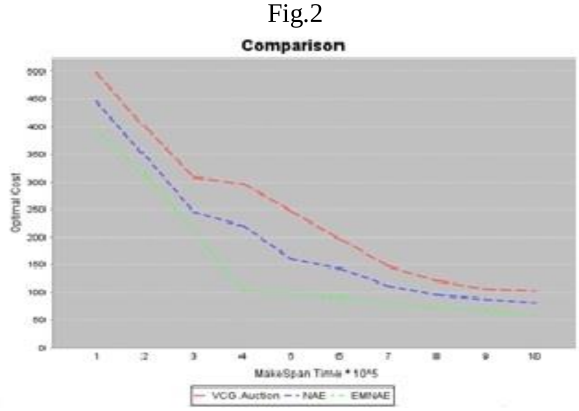


Fig.1 shows the output of execution time with respect to the size of job in Vickery-Clarke –Groves (VCG) Auction approach, Nash Auction Equilibrium (NAE) approach and the proposed Energy and job migration Nash Auction Equilibrium (EMNAE) approach. The proposed approach gives less execution time.

Fig.2 shows the result of Makespantime with respect to optimal cost in three approaches. In the proposed approach, the optimal cost value has been reduced.

Fig.3 shows the result of Migration Frequency with respect to number of jobs. The Proposed EMNAE approach has less migration frequency than the existing NAE approach.

Fig.4 shows the result of Energy versus number of jobs. The proposed EMNAE approach utilizes the low energy.

Fig.5 shows the output of Migration time with respect to number of jobs. In the proposed approach, the migration time is lesser than the other two existing approaches.

Fig.6 shows the output of Makespantime versus number of jobs. In proposed approach the Makespantime is lesser than the other two existing approaches.

Fig.3

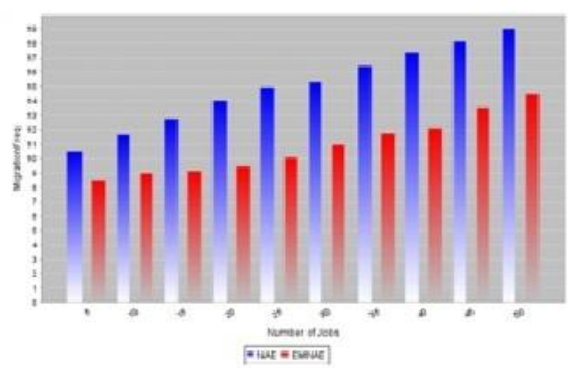


Fig.4

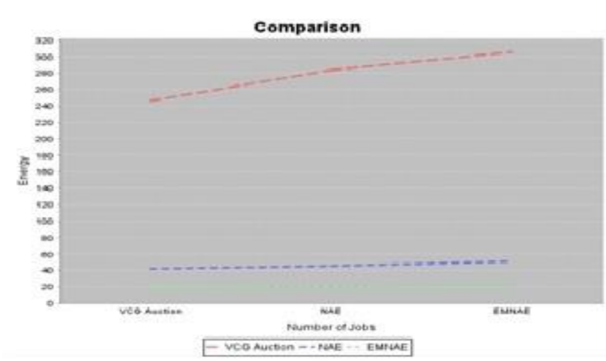


Fig. 5

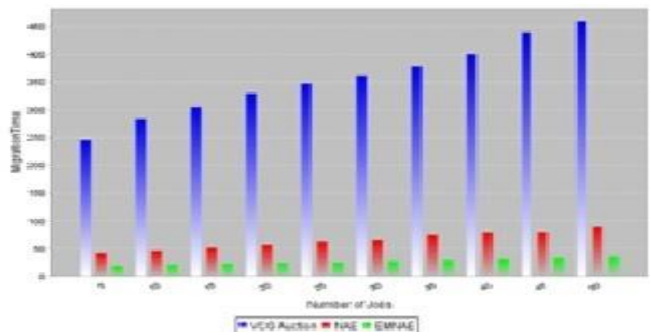
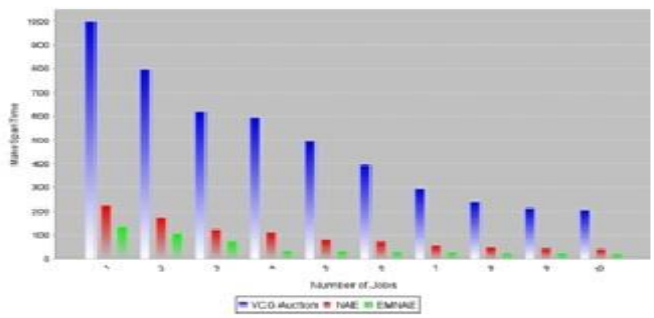


Fig. 6



V. Conclusion and Future Work

In the proposed strategy, optimal cost is estimated by considering load, Energy, jobs deadline, time to improve the efficiency in resource provisioning. In this proposed approach, Workflow logics, job migration, Fault tolerance services are used for resource provisioning in multi cloud environment. In future proposed system we introduce the new resource provisioning mechanism with the consideration of Load, Energy, and Network. By combining these three parameters we use the optimized fitness function for particle swarm optimization to select the computing node. Considering the energy consumption for the network transmission is to avoid the loss of energy. Nearby selection of the virtual machine is used to reduce the network transmission time as well as energy for network transmission. To select the optimal network path we use the Euclidian distance based node selection in future proposed system.

VI. REFERENCES

[1] Gideon Juve, Ewa Deelman "Resource Provisioning Options for Large Scale Scientific Workflows" in ieeeexplore.ieee.org e-science 2008.

[2] Mousumi Paul, Debabrata Samantha, Goutam Sanyal "Dynamic Job Scheduling in Cloud Computing based on Horizontal Load Balancing" in Int.J.comp.Tech.Appl Vol 2(5) 1552-1556.

[3] Raj Kumar Buyya, Anton Beloglazov and jemal Abawajy "Energy-Efficient Management of Data Center Resources for Cloud Computing: A Vision, Architectural Elements and Open Challenges" in www.cloudbus.org /papers/Greencloud2010.pdf

[4] Constanino Vazquez , Eduardo Haedo,Ruben.S.Montero "Dynamic Provision of Computing Resources from Grid Infrastructures and Cloud Providers" on Workshops at the Grid and Pervasive Computing conference 2009.

[5] K.C.Gowda,Radhika T.V,Akshatha.M "Priority based Resource Allocation Model for Cloud Computing" in International Journal of Science ,Engineering and Technology Research(IJSETR) Volume 2,Issue 1,January 2013.

[6] Stefan Schulte,Diecter Schuller,Philipp Hoenisch " Cost –Driven Optimization of Cloud Resource Allocation for elastic processes" in International Journal of Cloud Computing Vol.1 No.2 October-December 2013.

[7] Christina Hoffa, Gaurang Mehta, Timothy Freeman "On the Use of Cloud Computing for Scientific Workflows" in ieeeexplore.ieee.org escience 2008.

[8] Wenbing Zhao, D.M.Melliar-smith and L.E .Moser "Fault Tolerance Middleware for Cloud computing" on 2010 IEEE 3rd International conference on Cloud Computing.

[9] Md Imran Alam, Manjusha Pandey, Siddharth. S.Rautaray "A Proposal of Resource Allocation Management for Cloud Computing" in International Journal of Cloud Computing and Services Science (IJ-CLOSER) Vol., 3 No.2 April 2014, pp 79-86.

[10] Anju Bala, Inderveer chana "Fault Tolerance – Challenges, Techniques and Implementation in Cloud Computing" in International Journal of Computer Science Issues Vol .9, issue 1, No.1, January 2012.www.IJCSI.org.

[11] Awada Uchechukwu,Keqiu Li,Yanming Shen " Energy Consumption in Cloud Computing Data centers" in International Journal of Cloud Computing and services science Vol.3 No.3 June 2014 pp145-162 .

[12] Jing Liu,Xing Guo Luo "Job Scheduling Model for Cloud Computing Based on Multi Objective Genetic Algorithm" in International Journal of Computer Science Issues Vol.10 issue.1 No.3 January 2013 www.IJCSI.org.

[13] Nidhi Jain Kansal and Inderveer chana "Existing Load Balancing Techniques in Cloud Computing: A Systematic review" in Journal of Information Systems and communication Volume 3, issue 1, 2012, pp 87-91 http://www.bioinfo.in/contents.php?id=45

[14] M.Sudha , M.Monica "Investigation on Efficient Management of Workflows in Cloud Computing Environments" in International Journal on computer

Science and Engineering Vol.2 No.5,2010,1841-1845.
[15] Marcos Dias de Assuocao et al., "A Cost benefit analysis of using Cloud Computing to extend the capacity of Clusters" in Cluster comput (2010) 13: 335-347.
[16] Haiyang Qian and Deep Medhi "Server Operational Cost optimization for Cloud Computing Service Providers over a Time Horizon" in https:// www.usenix.org/event/notice11/tech/full-papers/Qian.pdf.
[17] Abirami.S.P and shalini Ramanathan " Linear Scheduling strategy for Resource Allocation in Cloud Environment" in International Journal on Cloud Computing Services and Architecture (IJCCSA) Vol.2 No.1 February 2012
[18] Mladen A.Vouk "Cloud Computing Issues, Research and Implementation" in Journal of Computing and Information Technology – CIT 16, 2008, 4,235-246.
[19] Amelie Chi Zhou and Bingsheng "Transformation-Based Monetary Cost Optimization for workflows in the Cloud "in IEEE Transactions on Cloud Computing Vol.2 No.1 January-March 2014.
[20] Martin Bichler, Jayant Kalagnanam "Configurable offers and winner determination in Multi attribute auctions "in Europeon Journal of Operational Research 160(2005) 380-394.